VARIABLES ALEATOIRES. LOIS DE PROBABILITE

I Définitions

On considère une expérience aléatoire dont l'univers est Ω.

Une **variable aléatoire** *X* est une fonction qui à chaque événement de Ω, associe un nombre réel.

Exemples : gain ou perte à un jeu, nombre de matches gagnés par une équipe, durée de vie d'une machine, temps d'attente à un guichet.

On dit que la variable aléatoire est **discrète** si elle ne prend que des valeurs isolées. Dans le cas contraire, on dit qu'elle est **continue**.

II Variables aléatoires discrètes

On se limitera au cas où le nombre de valeurs que peut prendre la variable aléatoire discrète X est **fini**.

1) Loi de probabilité

Si la variable aléatoire *X* prend *n* valeurs distinctes, on les notera x_i pour *i* variant de 1 à *n*.

On notera également p_i la probabilité que *X* prenne la valeur x_i.

La **loi de probabilité** de *X* est donnée par toutes les valeurs x_i et leur probabilité p_i. On la donne généralement sous la forme d'un tableau (où l'on peut vérifier que la somme de tous les p_i est égale à 1) :

x_i	x_1	x_2	....	x_n	Total
$P(X = x_i)$	p_1	p_2	....	p_n	1

2) Fonction de répartition

La **fonction de répartition** d'une variable aléatoire *X* est la fonction *F* qui, à tout nombre réel *x*, associe la probabilité de l'événement « $X \leq x$ »

$$\boxed{\text{Pour tout } x \in \mathbb{R}, \quad F(x) = P(X \leq x)}$$

On a alors $P(X > x) = 1 - F(X)$

- Cette fonction prend toutes ses valeurs dans l'intervalle [0 ; 1].
- Pour une variable aléatoire discrète, c'est une fonction en escalier.
- Elle vaut 0 sur l'intervalle $]-\infty\ ;\ x_1[$ et elle vaut 1 sur l'intervalle $[x_n\ ;\ +\infty[$.

3) Espérance et écart-type

L'**espérance mathématique** d'une variable aléatoire *X*, notée *E*, est la moyenne des valeurs prises par *X* pondérée par leur probabilité d'apparition.

$$\boxed{E(X) = \sum_{i=1}^{n} p_i x_i}$$

La **variance** de *X*, notée *V*, est l'espérance des carrés des écarts entre les valeurs prises par *X* et l'espérance de *X*.

$$V(X) = E\left[\left(X - E(X)\right)^2\right]$$

On démontre alors que $\boxed{V(X) = E(X^2) - \left(E(X)\right)^2}$

L'**écart-type** est la racine carrée de la variance. Il est noté σ et on a $\boxed{\sigma(X) = \sqrt{V(X)}}$

L'écart-type est un écart moyen par rapport à l'espérance : plus il est grand et plus les valeurs de *X* sont dispersées par rapport à cette espérance, c'est-à-dire plus la probabilité d'être proche de l'espérance est faible.

4) Exemple

On tire 5 cartes parmi un jeu de 32.

Si l'on n'obtient aucun as, on ne gagne rien. Si l'on obtient un as, on gagne 2 euros, si on obtient 2 as, on gagne 6 euros, si on tire 3 as, on gagne 12 euros et si l'on tire 4 as, on gagne 55 euros. La mise est de 2 euros. Le jeu est-il équitable ? Sinon, quelle doit être la mise pour qu'il le soit ?

Soit *X* la variable aléatoire égale au gain. Elle peut prendre les valeurs – 2, 0, 4, 10 ou 53

$$P(X=-2)=P(0\text{ as})=\frac{\binom{28}{5}}{\binom{32}{5}}=\frac{98280}{201376}\approx 0,4880 \;;\; P(X=0)=P(1\text{ as})=\frac{\binom{4}{1}\binom{28}{4}}{\binom{32}{5}}=\frac{81900}{201376}\approx 0,4067$$

$$P(X=4)=P(2\text{ as})=\frac{\binom{4}{2}\binom{28}{3}}{\binom{32}{5}}=\frac{19656}{201376}\approx 0,0976 \;;\; P(X=10)=P(3\text{ as})=\frac{\binom{4}{3}\binom{28}{2}}{\binom{32}{5}}=\frac{1512}{201376}\approx 0,0075$$

$$P(X=53)=P(4\text{ as})=\frac{\binom{4}{4}\binom{28}{1}}{\binom{32}{5}}=\frac{28}{201376}\approx 0,0002$$ d'où la loi de X est donnée par le tableau :

x_i	– 2	0	4	10	53	Total
$P(X = x_i)$	0,4880	0,4067	0,0976	0,0075	0,0002	1

On a alors $E(X) = 0{,}488 \times (-2) + 0{,}4067 \times 0 + 0{,}0976 \times 4 + 0{,}0075 \times 10 + 0{,}0002 \times 53 = -0{,}5$

Le gain moyen est de – 0,5 euros, donc le jeu n'est pas équitable, il est au désavantage du joueur. Pour qu'il le soit, il faudrait que le mise soit diminuée de 0,5 donc qu'elle soit de 1,50 euro.

Et $V(X)=0,488\times(-2)^2+0,4067\times0^2+0,0976\times4^2+0,0075\times10^2+0,0002\times53^2-(-0,5)^2=4,5754$ et $\sigma(X)\approx 2,14$

III Cas particulier : loi binomiale

1) Expérience de Bernoulli

On appelle **expérience de Bernoulli** une expérience aléatoire à deux issues possibles appelées succès S et échec $E=\overline{S}$.

On note p la probabilité du succès et q celle de l'échec. On a $q = 1 - p$

2) Schéma de Bernoulli

On appelle **schéma de Bernoulli** la répétition de n expériences de Bernoulli identiques et indépendantes.

Si p est la probabilité de chaque succès, on dit que le schéma de Bernoulli a pour **paramètres *n* et *p*.**

3) Loi binomiale

Étant donné un schéma de Bernoulli de paramètres n et p, on appelle X la variable aléatoire égale au **nombre de succès** obtenus au bout des n expériences.

X peut donc prendre les valeurs 0, 1, 2, ..., k, ..., n

On dit alors que la loi de probabilité de X est une **loi binomiale de paramètres *n* et *p***, et l'on a :

$$\boxed{\text{pour tout } k \text{ compris entre 0 et } n : P(X = k) = \binom{n}{k} p^k q^{n-k}}$$

On a alors $\boxed{E(X) = np}$ et $\boxed{V(X) = npq}$

IV <u>Variables aléatoires continues</u>

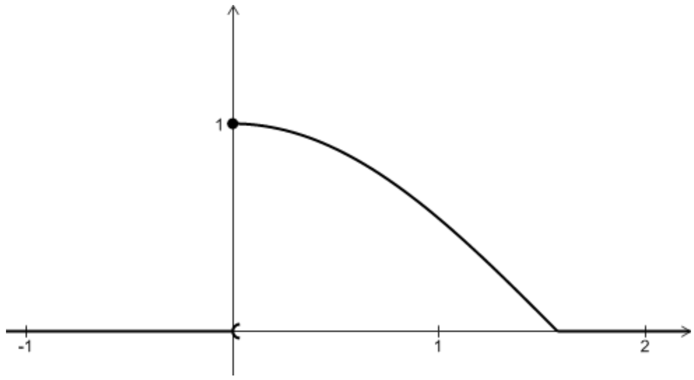
1) <u>Densité de probabilité</u>

Soit X une variable aléatoire continue. Donner la loi de probabilité de X, c'est donner une fonction f définie sur $\mathbb{R}$, appelée <u>densité de probabilité</u>, et qui vérifie :

- f est continue sur $\mathbb{R}$, sauf éventuellement en des points isolés
- f est positive sur $\mathbb{R}$
- $\int_{-\infty}^{+\infty} f(t)\,dt = 1$

exemple : Soit f la fonction définie sur $\mathbb{R}$ par :

$$\begin{cases} f(x) = 0 \text{ si } x < 0 \\ f(x) = \cos x \text{ si } x \in \left[0\,;\dfrac{\pi}{2}\right] \\ f(x) = 0 \text{ si } x > \dfrac{\pi}{2} \end{cases}$$



- f est continue sur $\mathbb{R}$ sauf en 0
- f est positive sur $\mathbb{R}$
- $\int_{-\infty}^{+\infty} f(t)\,dt = \int_{-\infty}^{0} f(t)\,dt + \int_{0}^{\frac{\pi}{2}} f(t)\,dt + \int_{\frac{\pi}{2}}^{+\infty} f(t)\,dt = \int_{-\infty}^{0} 0dt + \int_{0}^{\frac{\pi}{2}} \cos t dt + \int_{\frac{\pi}{2}}^{+\infty} 0dt = [\sin t]_0^{\frac{\pi}{2}} = \sin\dfrac{\pi}{2} - 0 = 1$

donc f est une densité de probabilité.

2) <u>Fonction de répartition</u>

Soit X une variable aléatoire continue et f sa densité de probabilité.

Alors la <u>fonction de répartition</u> F de X est la fonction définie sur $\mathbb{R}$ par $\boxed{F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt}$

Propriétés :

- $F' = f$
- F est continue car dérivable
- $\lim\limits_{x\to-\infty} F(x) = 0$ et $\lim\limits_{x\to+\infty} F(x) = 1$

Exemple : avec la densité de probabilité définie au paragraphe précédent, on a :

si $x \leq 0$, $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} 0dt = 0$

si $0 \leq x \leq \dfrac{\pi}{2}$, $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{0} f(t)\,dt + \int_{0}^{x} f(t)\,dt = \int_{-\infty}^{0} 0dt + \int_{0}^{x} \cos t dt = [\sin t]_0^x = \sin x$

si $x \geq \dfrac{\pi}{2}$, $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{0} f(t)\,dt + \int_{0}^{\frac{\pi}{2}} f(t)\,dt + \int_{\frac{\pi}{2}}^{x} f(t)\,dt = \int_{-\infty}^{0} 0dt + \int_{0}^{\frac{\pi}{2}} \sin t dt + \int_{\frac{\pi}{2}}^{x} 0dt = 0 + 1 + 0 = 1$

3) <u>Calculs de probabilités</u>

Théorème :

Soit X une variable aléatoire continue, f sa densité de probabilité et F sa fonction de répartition.

Soient a et b deux réels quelconques tels que $a \leq b$. Alors :

i) $P(X = a) = P(X = b) = 0$

ii) $P(X \leq b) = P(X < b) = F(b) = \int_{-\infty}^{b} f(t)\,dt$

iii) $P(X \geq a) = P(X > a) = 1 - F(a) = \int_{a}^{+\infty} f(t)\,dt$

iv) $P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b) = F(b) - F(a) = \int_{a}^{b} f(t)\,dt$

Démonstration : *i)* admis *ii)* Par définition de F, $P(X \le b) = F(b) = \int_{-\infty}^{b} f(t)\,dt$.

De plus, $P(X \le b) = P((X < b) \cup (X = b)) = P(X < b) + P(X = b) = P(X < b)$ d'après *i)*.

iii) D'après ce qui précède, $P(X \ge a) = P(X > a) = P(\overline{X \le a}) = 1 - P(X \le a) = 1 - F(a)$

De plus $1 - F(a) = \int_{-\infty}^{+\infty} f(t)\,dt - \int_{-\infty}^{a} f(t)\,dt = \int_{-\infty}^{+\infty} f(t)\,dt + \int_{a}^{-\infty} f(t)\,dt = \int_{a}^{+\infty} f(t)\,dt$

iv) Les quatre écritures sont équivalentes d'après *i)*, et l'on a :

$$P(X \le b) = P((X < a) \cup (a \le X \le b)) = P(X < a) + P(a \le X \le b)$$

d'où $P(a \le X \le b) = P(X \le b) - P(X < a) = F(b) - F(a) = \int_{-\infty}^{b} f(t)\,dt - \int_{-\infty}^{a} f(t)\,dt = \int_{a}^{b} f(t)\,dt$

V Lois particulières

1) Loi uniforme

Soit X une variable aléatoire continue.

X suit une loi uniforme sur l'intervalle $[a\,;b]$ si sa densité de probabilité f est définie par :

$$\begin{cases} f(x) = \dfrac{1}{b-a} \text{ si } x \in [a\,;b] \\ f(x) = 0 \text{ sinon} \end{cases}$$

Sa fonction de répartition est alors : $\begin{cases} F(x) = 0 \text{ si } x < a \\ F(x) = \dfrac{x-a}{b-a} \text{ si } x \in [a\,;b] \\ F(x) = 1 \text{ si } x > b \end{cases}$

En particulier si $a \le \alpha < \beta \le b$, on a $\boxed{P(\alpha \le X \le \beta) = \dfrac{\beta - \alpha}{b-a}}$

Exemple : On suppose que le temps d'attente en minutes à un arrêt de bus suit une loi uniforme sur l'intervalle [0 ; 12]

Dans ce cas, la probabilité d'attendre le bus entre 4 et 9 minutes est $P(4 \le X \le 9) = \dfrac{9-4}{12-0} = \dfrac{5}{12}$. C'est la même que celle d'attendre entre 0 et 5 minutes, ou qu'entre 6 et 11 minutes : c'est pour cette raison que la loi est dite uniforme.

2) Loi exponentielle

Soit X une variable aléatoire continue.

X suit une loi exponentielle de paramètre λ si sa densité de probabilité f est définie par :

$$\begin{cases} f(x) = \lambda e^{-\lambda x} \text{ si } x \ge 0 \\ f(x) = 0 \text{ sinon} \end{cases}$$

Sa fonction de répartition est alors : $\begin{cases} F(x) = 0 \text{ si } x < 0 \\ F(x) = 1 - e^{-\lambda x} \text{ si } x \ge 0 \end{cases}$

Exemple : On suppose que le temps d'attente à une caisse suit une loi exponentielle de paramètre $\lambda = 0{,}25$. La probabilité d'attendre plus de 4 minutes est donc $P(X \ge 4) = 1 - F(4) = 1 - \left(1 - e^{-0{,}25\times 4}\right) = e^{-1} \approx 0{,}36$

La probabilité d'attendre moins de 5 minutes est $P(X \le 5) = F(5) = 1 - e^{-0{,}25\times 5} = 1 - e^{-1.25} \approx 0{,}71$.

Remarque : Cette loi caractérise notamment les objets à « durée de vie sans vieillissement » : en effet, la probabilité que cet objet vive pendant une période donnée ne dépend que de cette durée, et non de l'âge de l'objet : autrement dit, l'espérance de vie ne dépend pas de l'âge ; ce peut être le cas de composants électroniques par exemple, mais hélas pas d'un être vivant.

Soit a et b deux réels tels que $0 < a < b$. On suppose que l'objet a atteint l'âge a et on cherche la probabilité pour qu'il atteigne (et donc dépasse) l'âge b : $P_{X \ge a}(X \ge b) = \dfrac{P((X \ge a) \cap (X \ge b))}{P(X \ge a)} = \dfrac{P(X \ge b)}{P(X \ge a)}$

d'où $P_{X \ge a}(X \ge b) = \dfrac{1 - F(b)}{1 - F(a)} = \dfrac{e^{-\lambda b}}{e^{-\lambda a}} = e^{-\lambda(b-a)}$ qui ne dépend que de la durée ($b - a$) et non de l'âge a.